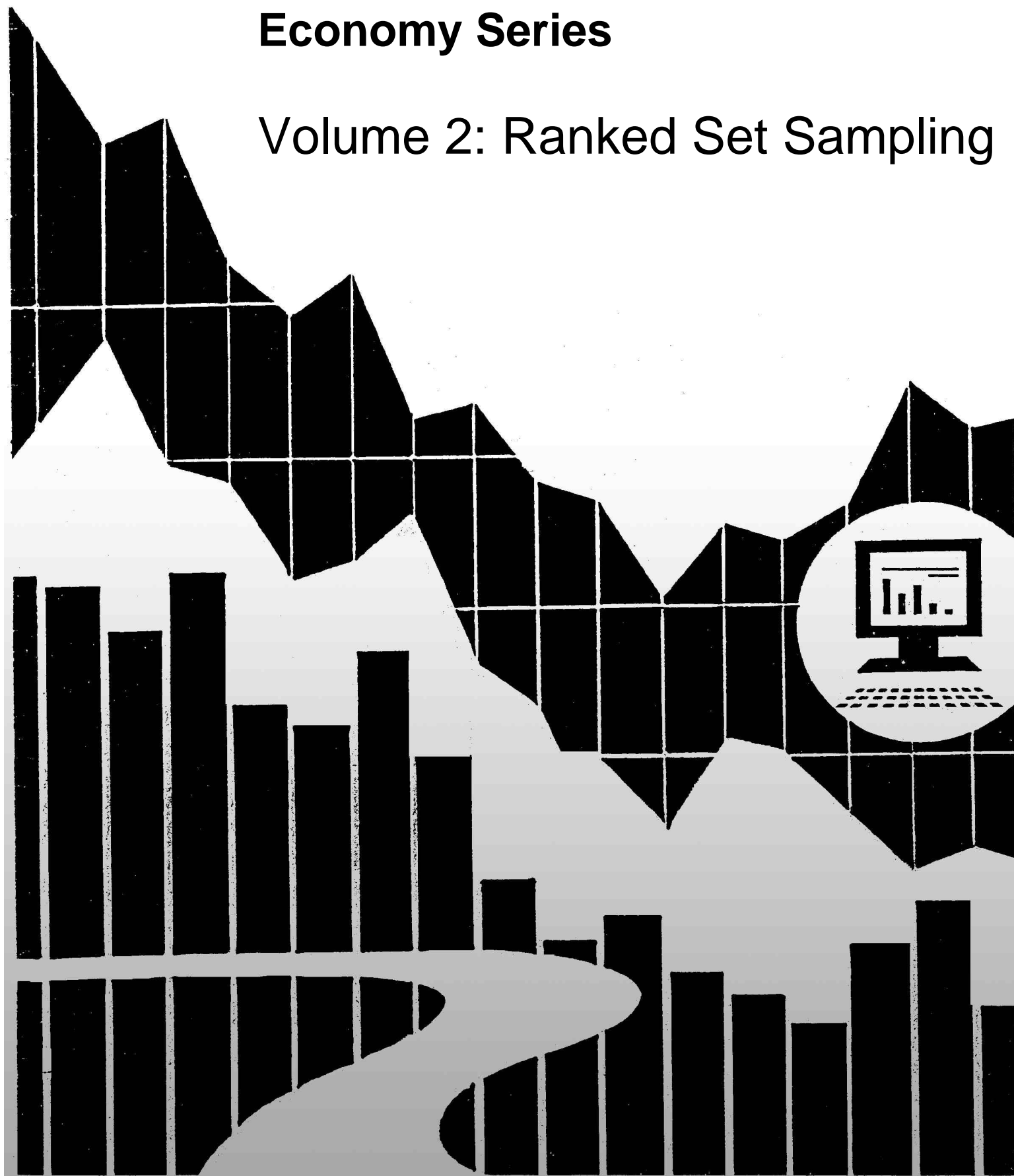




EPA Observational Economy Series

Volume 2: Ranked Set Sampling



EPA Observational Economy Series

Vol. 2: Ranked Set Sampling

United States
Environmental Protection
Agency

Policy, Planning,
And Evaluation
(2163)

EPA 230-R-95-006
August 1995

Contents

Foreword	iii
Acknowledgments	iv
1. Introduction	1
2. What is Ranked Set Sampling?	5
2.1. Description	5
2.1.1. Ranking Criteria	7
2.1.2. Robustness of the Procedure	7
2.2. Variations of the Basic Protocol	8
2.2.1. Unequal Allocation of Sample Units	8
2.2.2. Combining with Line Intercept Sampling	10
3. Applications	13
3.1. Forage Yields	13
3.2. Seedling Counts	14
3.3. Shrub Phytomass in Forest Stands	15
3.4. Herbage Mass	17
3.5. PCB Contamination Levels	18
3.6. Improved Compositing of Samples	21
3.7. Additional Applications	22
4. Summary	25
References	27

Foreword

Ranked set sampling is a novel method for achieving observational economy when performing environmental monitoring and assessment. Compared to simple random sampling, ranked set sampling yields a sample of observations that are more representative of the underlying population. Therefore, either greater confidence is gained for a fixed number of observations, or for a desired level of confidence, less observations (less \$) are needed.

The increased sampling efficiency is achieved by exploiting auxiliary information involving acquired field samples, a characteristic of double sampling procedures. With ranked set sampling, however, the auxiliary information does not have to be a quantitative concomitant variable. In fact, it can be purely judgmental; and thus, in the spirit of total quality management, it stimulates and utilizes a productive cross disciplinary dialogue among those responsible for sampling and assessment. Additionally, the ranked set sampling procedure is robust in the sense that it cannot perform worse than the usual simple random sampling.

This volume in the EPA Observational Economy Series introduces the concept and method of ranked set sampling for its timely inclusion in the toolbox of sampling procedures that aim to achieve observational economy, particularly when analytical costs dominate the monitoring scenario.

Acknowledgments

The EPA Observational Economy Series is a result of the research conducted under a cooperative agreement between the U.S. Environmental Protection Agency and the Pennsylvania State University Center for Statistical Ecology and Environmental Statistics, Professor G.P. Patil, Director.

The EPA Grant CR-821531010, entitled "Research and Outreach on Observational Economy, Environmental Sampling and Statistical Decision Making in Statistical Ecology and Environmental Statistics" consists of ten separate projects in progress at the Penn State Center: 1) Composite Sampling and Designs; 2) Ranked Set Sampling and Designs; 3) Environmental Site Characterization and Evaluation; 4) Encounter Sampling; 5) Spatio-temporal Data Analysis; 6) Biodiversity Analysis and Monitoring; 7) Adaptive Sampling Designs; 8) Statistics in Environmental Policy and Regulation for Compliance and Enforcement; 9) Statistical Ecology and Ecological Risk Assessment; and 10) Environmental Statistics Knowledge Transfer, Outreach and Training.

The series is published by the Statistical Analysis and Computing Branch of the Environmental Statistics and Information Division in the EPA Office of Policy, Planning and Evaluation. This volume in the series is largely based on the work of G. D. Johnson, A. Kaur, G. P. Patil, A. K. Sinha and C. Taillie at the Penn State Center in cooperation with John Fritzvold, Herbert Lacayo, Robert O'Brien, Brenda Odom, Barry Nussbaum, and John Warren, as project officers at the U.S. EPA. Questions or comments on this publication should be directed to Dr. N. Phillip Ross, Director, Environmental Statistics and Information Division (Mail Code 2163), United States Environmental Protection Agency, 401 M Street SW, Washington, DC 20460; Ph. (202) 260-2680.

1. Introduction

Environmental monitoring and assessment mostly requires observational data, as opposed to data obtained from controlled experiments. This is true whether we are assessing the extent of soil contamination at a one-acre site or some measure of forest resources over the Pacific Northwest region of the United States. Obtaining such data requires identification of sample units to represent the population of concern, followed by selection of particular units to quantify the characteristic(s) of interest. Sample units are basically the smallest units of measurement such as plots, soil cores, individuals, etc., while typical characteristics of interest include biomass, chemical concentrations or “head counts”.

Typically the most expensive part of this process is laboratory analysis, while identification of potential sample units is a comparatively simple matter. We can therefore achieve great observational economy if we are able to identify a large number of sample units to represent the population of interest, yet only have to quantify a carefully selected subsample.

This potential for observational economy was recognized for estimating mean pasture and forage yields in the early 1950’s when McIntyre (1952) proposed a method, later coined Ranked Set sampling (RSS) by Halls and Dell (1966), and currently under active investigation in various quarters.

As a simple introduction to the concept of RSS, consider the following example:

Let’s say we wish to estimate the mean height of students at a university from a random sample of three students. Furthermore, in order to acknowledge the inherent uncertainty, we need to present this estimate as a confidence interval within which we expect the true population mean to lie with desired confidence.

Now the simplest way to obtain our sample is to randomly select three students from the university’s population, then measure their heights. While the arithmetic average of the three heights is an unbiased point estimate of the population mean, the associated confidence interval can be very large, reflecting the high degree of uncertainty with estimating a large population mean with only three measurements. This is because we have no control over

which individuals of the population enter the sample. For example, we may happen to grab two very short people and one very tall; or we may grab three very tall people. The only way to overcome such a problem with a simple random sample (SRS) is to increase the sample size.

On the other hand, we may obtain a ranked set sample. To do this, we may randomly invite three students to breakfast and visually rank them with respect to height. We then select the student we believe is shortest and actually measure his or her height. Repeating this process with lunch, we then select the middle ranked person, and, as such, select the tallest ranked person at dinner. The resulting measurements of student heights constitute a ranked set sample. As with the SRS measurements, the arithmetic average of the RSS measurements provides an unbiased point estimate of the population mean; however, the associated confidence interval can potentially be much smaller than that obtained with SRS measurements, thus reflecting decreased uncertainty. This encouraging feature results because measurements obtained through RSS are likely to be more regularly spaced than those obtained through SRS and therefore are more representative of the population. Amazingly, the RSS procedure induces stratification of the whole population at the sample level; in effect, we are randomly sampling from the subpopulations of predominantly short, medium and tall students without having to construct the subpopulation strata. Each subpopulation has its own distribution, as visualized in Figure 1, where we see how the parent population gets effectively partitioned into subpopulations.

McIntyre's proposal does not appear to have been applied for over a decade, after which forest and range researchers continued to discover the effectiveness of RSS (see Halls and Dell, 1966; Evans, 1967; Martin, et al. 1980; Jewiss, 1981; and Cobby, et al. 1985). Theoretical investigations by Dell and Clutter (1972) showed that, regardless of ranking errors, the RSS estimator of a population mean is unbiased and at least as precise as the SRS estimator with the same number of quantifications. David and Levine (1972) investigated the case where ranking is done by a numerical covariate.

Furthermore, RSS also provides more precise estimators of the variance (Stokes, 1980a), the cumulative distribution function (Stokes and Sager, 1988) and at times the Pearson correlation coefficient (Stokes, 1980b). For an annotated bibliography with an historic perspective, see Kaur, Patil, Sinha, and Taillie (1995).

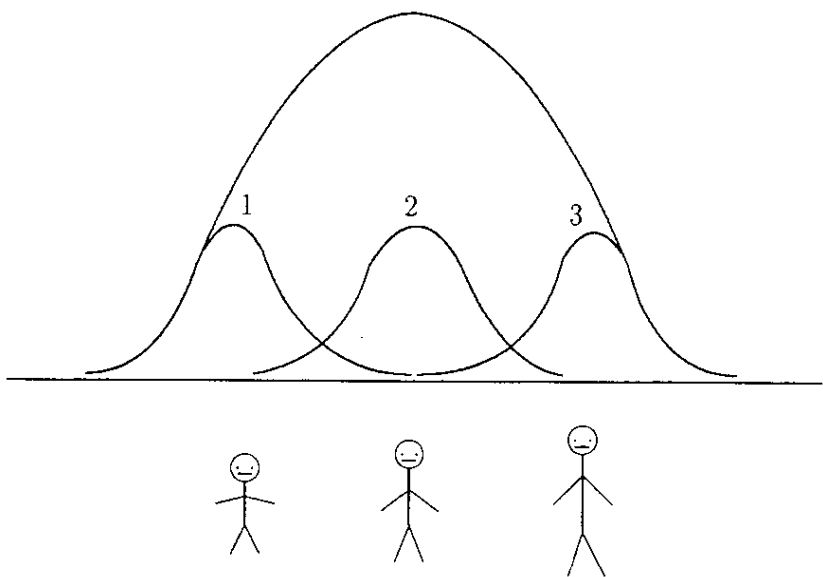


Figure 1: Frequency distributions of heights of different ranks superimposed on population frequency distribution of all heights—a schematic diagram.

2. What is Ranked Set Sampling?

2.1. Description

As mentioned in Chapter 1, to create ranked sets we must partition the selected first phase sample into sets of equal size. In order to plan an RSS design, we must therefore choose a set size which is typically small, around 3 or 4, to minimize ranking error. Let's arbitrarily call this set size m , where m is the number of sample units allocated into each set. Now proceed as follows.

- step 1: Randomly select m^2 sample units from the population.
- step 2: Allocate the m^2 selected units as randomly as possible into m sets, each of size m .
- step 3: Without yet knowing any values for the variable of interest, rank the units within each set based on a perception of relative values for this variable. This may be based on personal judgment or done with measurements of a covariate that is correlated with the variable of interest.
- step 4: Choose a sample for actual analysis by including the smallest ranked unit in the first set, then the second smallest ranked unit in the second set, continuing in this fashion until the largest ranked unit is selected in the last set.
- step 5: Repeat steps 1 through 4 for r cycles until the desired sample size, $n = mr$, is obtained for analysis.

As an illustration, consider the set size $m = 3$ with $r = 4$ cycles. This situation is illustrated in Figure 2 where each row denotes a judgment-ordered sample within a cycle, and the units selected for quantitative analysis are

cycle	rank		
	1	2	3
1	⊙	.	.
	.	⊙	.
	.	.	⊙
2	⊙	.	.
	.	⊙	.
	.	.	⊙
3	⊙	.	.
	.	⊙	.
	.	.	⊙
4	⊙	.	.
	.	⊙	.
	.	.	⊙

Figure 2: A ranked set sample design with set size $m = 3$ and the number of sampling cycles $r = 4$. Although 36 sample units have been selected from the population, only the 12 circled units are actually included in the final sample for quantitative analysis.

circled. Note that 36 units have been randomly selected in 4 cycles; however, only 12 units are actually analyzed to obtain the ranked set sample of measurements.

Obtaining a sample in this manner results in maintaining the unbiasedness of simple random sampling; however, by incorporating “outside” information about the sample units, we are able to contribute a structure to the sample that increases its representativeness of the true underlying population.

If we quantified the same number of sample units, $mr = 12$, by a simple random sample, we have no control over which units enter the sample. Perhaps all the 12 units would come from the lower end of the range, or perhaps most would be clustered at the low end while one or two units would come from the middle or upper range. With simple random sampling, the only way to increase the prospect of covering the full range of possible values is to increase the sample size. With ranked set sampling, however, we increase the representativeness with a fixed number of sample units, thus saving considerably on quantification costs.

With the ranked set sample thus obtained, it can be shown that unbiased estimators of several important population parameters can be calculated, including the mean and, in case of more than one sampling cycle, the variance.

2.1.1. Ranking Criteria

A real key to success lies with step 3 in the above procedure—ranking. This may be based on visual inspection or other expert opinion about the sample units. For example, a field-seasoned range scientist or woods person may readily be able to rank three or four quadrats of grass with respect to overall volume or mass. Meanwhile a hazardous waste site inspector may be able to reliably rank areas of soil with respect to concentrations of a toxic contaminant, based on features like surface staining, discoloration or the appearance of stressed vegetation.

On the other hand, if another characteristic is available that is highly correlated with the characteristic of interest but costs much less to obtain, then we may rank by the values of such a “covariate”. For example, reflectance intensity of near-infrared electromagnetic radiation, as recorded in a remotely sensed digital image, is directly proportional to vegetation concentration on the ground. Another example might be to measure total organic halides (TOX) in soil in order to rank soil sampling units with respect to the concentration of volatile organic solvents. As an indicator variable, TOX is much less expensive to measure than specific organic compounds.

2.1.2. Robustness of the Procedure

Several questions may now arise, such as:

1. What if the distribution of sample measurements is skewed? or symmetric? or essentially unknown?
2. What if the sample units are not randomly allocated into sets?
3. How does error in ranking affect results?

First of all, while independent (random) and identically distributed sample measurements obtained through perfect ranking may lead to *optimum* performance of ranked set sampling, no matter how much these desirable characteristics are deviated from, the sampling efficiency will never be worse than with simple random sampling using the same number of quantifications. In fact, when efficiency is expressed as the relative precision (RP) such that

$$RP = \frac{\text{variance of sample average with simple random sampling}}{\text{variance of sample average with ranked set sampling}},$$

it can be shown that the bounds of this relative precision are

$$1 \leq RP \leq (m+1)/2,$$

where m is the set size. Since RP can not be less than one, the RSS protocol can not be worse than the simple random sampling protocol.

2.2. Variations of the Basic Protocol

2.2.1. Unequal Allocation of Sample Units

The performance of RSS decreases as the underlying distribution of the characteristic of interest becomes increasingly skewed. McIntyre (1952) originally suggested that this problem may be overcome by allocating sample units into ranks in proportion to the standard deviation of each rank. This is the same approach as used in stratified random sampling, known as Neyman allocation, and would indeed be optimal if we had reliable prior estimates of the rank standard deviations. An example of unequal allocation is displayed in Figure 3. Here we have the same set size, $m = 3$, and sample size, $n = 12$, as in the earlier example of equal allocation; however, the number of sampling cycles is adjusted so as to yield the desired unequal allocation of samples.

Unequal allocation can actually increase the performance of RSS above and beyond that achievable with standard equal allocation; however, if not properly applied, the performance of RSS can be worse than the performance of simple random sampling. Actually, the bounds on relative precision with unequal allocation become

$$0 \leq RP \leq m.$$

indicating that, with appropriate unequal allocation, the relative precision may even increase to a level of m , and not just $(m + 1)/2$ as in the case with equal allocation.

Although an optimal RSS design would allocate samples into ranks in direct proportion to the rank standard deviations, we rarely know the standard deviations beforehand. We do know, however, that the distributions of many environmental and ecological variables are skewed towards the right, meaning that while most values are clustered around a median, a few much larger values are usually present. This skewness can actually be exploited to increase the precision beyond that obtained with ranked set sampling under equal allocation because standard deviations usually tend to increase with increasing rank values for right-skewed distributions. With some idea of the degree of skewness, Kaur, Patil, and Taillie (1994) have devised a rule-of-thumb for allocating sample units into ranks that performs closely to the optimal Neyman allocation. Therefore, distributions of many environmental and ecological variables may actually lend themselves well to being estimated

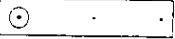
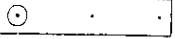
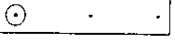
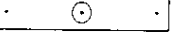
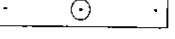
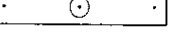
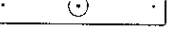
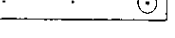
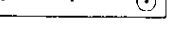
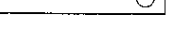
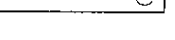

Sets	Units	No. of sets
1		3
2		
3		
4		4
5		
6		
7		
8		5
9		
10		
11		
12		

Figure 3: Ranked set sampling with unequal allocation: circles indicate sample units chosen for quantification.

with very high precision relative to that obtainable through simple random sampling.

2.2.2. Combining with Line Intercept Sampling

A common field sampling method for ecological assessments is to include sample units that one encounters along a line (transect) that is randomly selected within a two-dimensional area of interest. Units are typically members of a plant or animal species.

Often the number of sample units identified are too numerous to select every one for quantification, especially if measurements are destructive, such as with cutting vegetation for weighing. If the initially identified sample units are treated as a larger first phase sample, $n' = m^2r$, then the RSS protocol can be applied to select a smaller subsample, $n = mr$, for actual quantification. For example, consider a single sampling cycle when the set size, m equals three for estimating the biomass of shrubs in a given area. A line transect for such a situation may be visualized as in Figure 4.

Such an RSS-based line intercept sample has been found to produce more precise, and still unbiased, estimators of the population mean, size, total and cover, compared to the SRS-based line intercept sample (see Muttlack and McDonald, 1992).

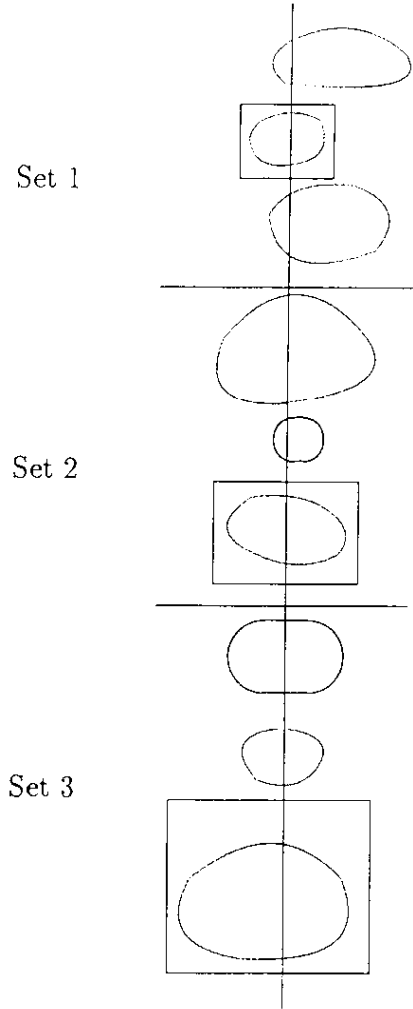


Figure 4: Aerial view of a line transect intercepting shrubs. For set size $m = 3$, nine shrubs are partitioned into 3 sets of 3. Using apparent shrub size for ranking with respect to biomass, the shrubs taken for analyses include the smallest ranked in the first set, the second smallest ranked in the second set and the largest ranked in the third set.

3. Applications

3.1. Forage Yields

Although McIntyre's original proposal of estimating pasture yields by "unbiased selective sampling using ranked sets" was made in 1952, no applications were apparently reported until fourteen years later. Halls and Dell (1966) applied McIntyre's method, coining it "ranked set sampling" for estimating the weights of browse and herbage in a pine-hardwood forest of east Texas. These authors discovered RSS to be considerably more efficient than SRS.

Sets of three closely grouped quadrats were formed on a 300-acre tract. At select locations, metal frames of 3.1 square feet were placed at three randomly selected points within a circle of 13 foot radius as seen in Figure 5. Quadrats were then ranked as lowest, intermediate and highest according to the perceived weight of browse and, separately, of herbage. Then, after clipping and drying, the separate weights of browse and herbage were determined for each quadrat. This was repeated for 126 sets for estimating browse and 124 sets for estimating herbage.

In order to simulate the SRS estimator for the mean weight of browse, one quadrat was randomly selected from each set without considering its rank. Since actual values were known for each quadrat, the RSS estimator was obtained by randomly choosing the ranks to be quantified for each set, resulting in 37 lowest ranks, 46 intermediate ranks and 43 highest ranks. Halls and Dell also examined McIntyre's suggestion that unequal allocation might further improve the efficiency of estimation. Since the standard deviations for the order statistics were 7, 13 and 27.7 for the low, intermediate and high yield, respectively (ratio of 1:2:4), they selected 14 quadrats in the low group, 40 in the intermediate group and 72 in the high group. Note that perfect ranking was obtained for both RSS protocols because the actual values were already known for each quadrat.

Results of these three sampling protocols are reported in Table 1. As expected under perfect ranking, precision due to RSS with approximately equal allocation increased, more than doubling for browse estimates. Furthermore, when allocation was proportional to the order statistic standard deviation,

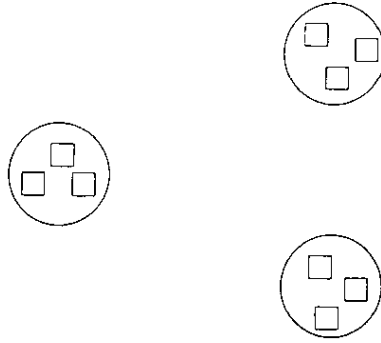


Figure 5: Within each circle, quadrats are randomly placed, followed by ranking and analysis of one appropriate quadrat. (not to scale)

Table 1: Summary statistics for browse and herbage estimates

	<u>browse</u>		<u>herbage</u>	
		Variance		Variance
	mean	of mean	mean	of mean
Unranked: random	14.9	4.55	7.3	1.00
Perfect ranking:				
near equal allocation	13.2	2.18	7.0	.73
Perfect ranking:				
proportional allocation	12.9	1.91	7.2	.58

(Source: Halls and Dell. 1966)

the precision increased still further, thus supporting McIntyre's contention.

Another very valuable aspect of this study was that two observers independently ranked the quadrats, one a professional range man and the other a woods worker. There was practically no difference in the ranking results between the two observers.

3.2. Seedling Counts

The effectiveness of RSS for improving the sampling precision of seedling counts was studied by Evans (1967) in an area in central Louisiana that was seeded to Longleaf Pine (*Pinus palustris mill.*). After dividing the target area into 24 blocks, each block was then subdivided into 25 one-milacre plots. All

600 plots were initially measured to characterize the population, which is summarized in Table 2a. The population mean and standard deviation were calculated to be 1.675 and 1.36, respectively.

For the RSS protocol, three plots were randomly selected from each of the 24 blocks (sets), resulting in 72 identified plots. The three plots within each set were then visually ranked. One cycle consisted of selecting the lowest ranked plot from the first set, the second lowest from the second set and the highest ranked plot from the third set. Repeating the cycle eight times yielded 24 selected plots in the ranked set sample ($m = 3, r = 8$). This whole procedure was repeated twice so that three separate field trials were performed, as summarized in Table 2b. Evans also computed the means and standard deviations of each rank using all 72 identified plots for each of the three field trials. These results are reproduced in Table 2c for comparison to the RSS results in Table 2b.

In order to compare RSS to SRS, Evans resampled the 24 blocks (sets) 80 times to obtain two empirical distributions of the means, one based on the RSS estimator and the other based on the SRS estimator, which is actually a stratified random sample estimator. The results of this “bootstrapping” exercise are reproduced in Table 2d where we see a significant reduction in the variance due to RSS.

3.3. Shrub Phytomass in Forest Stands

The performance of RSS for estimating shrub phytomass (all vegetation between one and five meters high) was evaluated by Martin et al. (1980) at a forested site in Virginia. They investigated four major vegetation types along a decreasing moisture gradient: mixed hardwood, mixed oak, mixed oak and pine, and mixed pine. For each vegetation type, a 20m by 20m area was subjectively located which was further divided into 16 plots of equal size (5m by 5m).

For the RSS procedure, four sets of four plots were randomly selected from the 16 plots in each vegetation type. The plots in each set were then ranked by visual inspection, followed by quantifying the smallest ranked plot from the first set, the second smallest ranked plot from the second set and so on in the usual manner for RSS. This was repeated for each of the four vegetation types. For the SRS procedure, four out of the 16 plots in each vegetation type were randomly selected without replacement, followed by quantification of each selected plot. Again, this is actually a stratified random sample since each vegetation type is a separate stratum. Shrub phytomass was also determined for all 64 plots to obtain a grand mean and variance for comparison. Their results are reproduced in Table 3 where we see a

Table 2: Data from Longleaf Pine Seedling Counts

(a) The frequency distribution of seedling counts in the 600 milacre plots.

Seedling Count	0	1	2	3	4	5	6	7	8	9
Frequency	110	201	157	75	33	17	3	3	0	1

(b) Means and variances of three ranked set sample trials. ($mr = 24$)

Trial	Mean	Variance
1	1.49	0.043
2	1.62	0.056
3	1.71	0.024

(c) Means and standard deviations of all seedlings for all ranks of three field trials of ranked set sampling.

Trial	Means			Mean	Standard Deviations		
	L	M	H		L	M	H
1	0.750	1.500	2.625	1.625	0.532	0.750	1.173
2	0.917	1.625	2.833	1.792	0.881	1.013	1.880
3	0.750	1.708	3.125	1.861	0.520	0.955	0.927

(d) Test of significance of ranked-set versus random sampling.

Method of sampling	Number applications	Degrees freedom	Mean	Sum squares	Variance	F
Random	80	79	1.709	7.572	.0958	3.91**
Ranked-set	80	79	1.647	1.939	.0245	

**Significant at the .01 level of probability

(Source: Evans, 1967)

Table 3: RSS and SRS results for 16 measured plots across all vegetation types.

Sampling Method	Mean Phytomass (kg/ha)	Variance of the Mean ($\times 10^6$)	Coefficient of Variation of the Mean(%)
All 64 Plots	2536	0.15	15
SRS	1976	4.54	108
RSS	2356	2.73	70

(Source: Martin et al. 1980)

substantial increase in precision of the mean estimator associated with RSS.

3.4. Herbage Mass

In order to compare RSS to SRS for estimating herbage mass in pure grass swards and both herbage mass and clover content in mixed grass-clover swards, Cobby et al. (1985) conducted four experiments at Hurley (UK). Besides comparison of RSS to SRS, their objective was to assess the effects of the following factors on RSS: (i) imperfect ranking within sets, (ii) greater variation between sets than within sets, and (iii) asymmetric distribution of the quantified values.

The first two experiments were conducted by randomly selecting 15 locations, followed by randomly selecting three quadrats at each location and having several observers rank the quadrats within each set. For the last two experiments, 45 quadrats were drawn at random from the entire target area. This allowed an assessment of the effects of both spatial variation and ranking errors within sets.

Their results are reproduced in Table 4, where RP of both the worst and best observers are compared to the RP under perfect ranking, and the between and within set variances are presented for assessing spatial variation. These authors determined the main adverse factor to be within set clustering, and they recommend spacing quadrats within sets as far apart as possible when local spatial autocorrelation exists. With this in mind, they recommend RSS over SRS for sampling grass and grass-clover swards.

Table 4: Relative precisions (RP) \pm s.e. of the worst and the best observers, and under perfect ranking; and the between and the within set variances while estimating herbage mass (grass and mixture) and clover contents.

Experiments	Relative Precisions (R P)			Variances	
	Worst	Best	Perfect	Between	Within
1 (Grass)	1.11 \pm 0.09	1.23 \pm 0.14	1.31 \pm 0.17	0.24	0.31
2 (Mixture)	1.11 \pm 0.09	1.27 \pm 0.10	1.40 \pm 0.16	0.07	0.09
3 (Grass)	————	————	1.66 \pm 0.17	0.00	1.58
4 (Mixture)	1.36 \pm 0.14	1.51 \pm 0.15	1.55 \pm 0.16	0.11	0.66
2 (Clover)	1.15 \pm 0.12	1.34 \pm 0.15	1.44 \pm 0.16	16.3	34.4
4 (Clover)	1.36 \pm 0.19	1.62 \pm 0.18	1.72 \pm 0.20	16.2	71.6

(Source: Cobby et al. 1985)

3.5. PCB Contamination Levels

Before being lead to believe that RSS is only for vegetation studies, let us consider estimating PCB concentrations in soil. Patil, Sinha, and Taillie (1994) used measurements of this contaminant collected at a Pennsylvania site along the gas pipeline of the Texas Eastern Company. Table 5 provides the summary statistics of PCB values in two sampling grids (A and C) within this site. Since the distribution of these data was highly skewed, they examined the effects of unequal as well as equal allocation of samples. More specifically, they examined the following schemes:

- (a) Equal allocation of samples using all possible choices of sample units of each set size,
- (b) Equal allocation of samples for a particular sample, and
- (c) Unequal allocation of samples.

Considering set sizes 2, 3, and 4, the relative savings (RS) were computed as $\frac{\text{var}(\text{SRS}) - \text{var}(\text{RSS})}{\text{var}(\text{SRS})}$ taking into consideration all possible choices of sample units for each set size for both the grids under the equal allocation scheme. The results are given in Table 6, where it is evident that RS increases with set size but that the magnitude of RS is higher for grid C than for grid A. Note that the data for grid C is much less skewed than grid A, as seen in Table 5.

Table 5: Descriptive statistics of PCB values in grids A and C.

Characteristics	Grid	
	A	C
Number of Observations	184	68
Mean	200.9	600.2
Standard Deviation	902.9	1585
Coefficient of Variation	4.49	2.64
Coefficient of skewness	9.27	4.48
Coefficient of kurtosis	99.69	20.88

Table 6: Relative savings (RS) considering all possible combinations of each set size under perfect ranking situation with equal allocation.

Set Size (m)	Grid	
	A RS	C RS
2	4	9
3	7	16
4	10	22

Table 7: Values of the sample mean, $\bar{X}_{(m)u}$, relative precision, and relative savings under the perfect ranking protocol with unequal allocation of samples.

Set Size m	Proportion of samples (Exact No.)	$\bar{X}_{(m)u}$	A			Grid			C
			RP	RS	Proportion of samples (Exact No.)	$\bar{X}_{(m)u}$	RP	RS	
2	1:10 (8,84)	205.9	1.724	42	1:10 (3,31)	535.2	2.041	51	
2	1:15 (6,86)	203.1	1.818	45	1:15 (2,32)	520.4	2.174	54	
3	1:4:20 (2,10,48)	203.6	2.174	54	1:1.7:1.5 (5,8,8)	560.1	1.471	32	
3	1:4:25 (2,8,50)	201.1	2.326	57	1:2:7 (2,4,15)	615.2	1.923	48	
4	1:3:5:16 (2,5,9,28)	247.1	1.695	41	1:2:3:4 (2,3,5,6)	576.6	2.083	52	
4	1:3:9:27 (2,2,10,30)	226.1	1.316	24	1:1:3:5 (2,2,4,8)	802.4	1.449	31	

For comparing the performance of the RSS protocol relative to that of SRS with unequal allocation of samples, these authors considered two different proportional allocations for each set-size in order to decide the sample size for each rank. This has been done to show the impact of proportional allocation on the magnitude of relative savings accrued due to RSS over SRS. The results are given in Table 7, where the magnitudes of relative savings are seen to be quite substantial for each set size for both the grids.

While unequal allocation of samples into ranks can substantially increase RS when the underlying population follows a skewed distribution, this procedure does require some prior knowledge of the underlying distribution. For this purpose one may either take advantage of prior surveys of similar nature or conduct a pilot study. This same problem also arises in determining the optimum sample size under Neyman's allocation scheme for stratified random sampling. Recent work by Kaur, Patil and Taillie (1994) has addressed the issue of optimum allocation when some knowledge about the underlying distribution is available, and they have devised a rule-of-thumb for allocating sample units based on skewness.

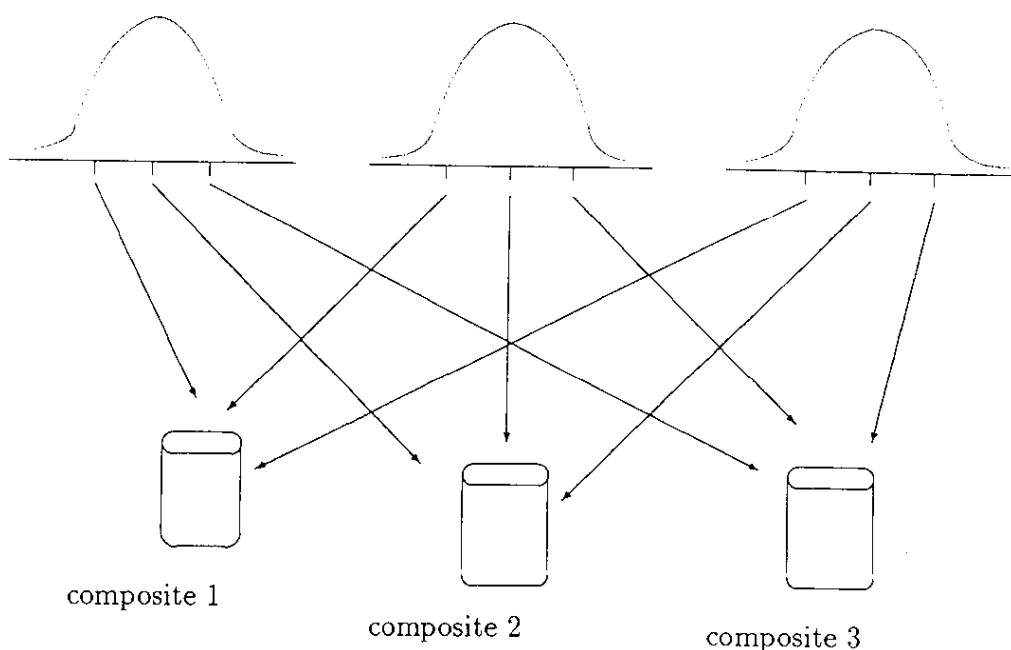


Figure 6: Formation of three composites from three ranked set samples, each with a set size of three. Homogeneity is maximized (variability is minimized) within each composite by forming composites from equally ranked samples.

3.6. Improved Compositing of Samples

Consider a situation that calls for composite sampling, as discussed in Volume 1 of this series. If our primary objective is classification of the individual samples used to form the composites and/or identification of those individual samples that constitute an upper percentile with respect to the characteristic of interest, then we will need to retest certain individual samples. Since the purpose of composite sampling is to minimize the number of analytical tests required, we obviously want to minimize the extent of retesting individual samples. Maximizing the homogeneity within the composites will minimize the necessary number of retests. Therefore, it is desired to form composites from individual sample units that are as much alike as possible.

As pointed out by Patil, Sinha, and Taillie (1994), one can increase the chances of obtaining maximum homogeneity within composites by forming composites from samples identified to be in the same rank as conceptualized in Figure 6. The RSS protocol can thus be combined with composite sampling to achieve even greater observational economy than composite sampling alone.

3.7. Additional Applications

Yanagawa and Chen (1980) mention that the RSS technique is regularly employed at the Pastoral Research Laboratory, CSIRO at Armidale, N.S.W., Australia. A plate with four holes is randomly thrown on a field and the pasture in each hole is ranked by eye, followed by selection of one hole for quantification of pasture. These authors also mention that RSS has been used to estimate rice crops in Okinawa, Japan. They attribute this information to H. Mizuno at the "Mathematical Method in Sampling" symposium held at Chiba University (Japan), 1974.

In addition to the reported applications of RSS, several other applications have been recommended:

- (i) Evans (1967) pointed out that the method would prove time-saving in the determination of cell wall thickness of different species of wood. In the same area of application, Dell (1969) has mentioned that the RSS procedure should be efficient for estimating averages for various properties of cells in a cross section of wood chips.
- (ii) The RSS method of sampling could be useful in determining the average length of various kinds of bacterial cells. Also, it may be used to determine the average number of bacterial cells per unit volume. This is possible because it is convenient to order test tubes containing the cell suspension on the basis of concentration with the help of an optical instrument without knowing the exact number of the bacterial cells. Takahasi and Wakimoto (1968) have suggested these applications.
- (iii) The technique may also be used to determine the average height of trees because it is easy to rank the heights of several nearby trees by a visual inspection. This application has also been mentioned by Takahasi and Wakimoto (1968).
- (iv) Stokes and Sager (1988) have suggested that the method of RSS could also be used to investigate a difficult-to-measure characteristic in human populations. They have, for example, referred to the Consumer Expenditure Survey. The results of this survey are used for the construction of the Consumer Price Index. But this survey requires detailed record keeping by the participating households as well the services of professional interviewers. In this situation a pre-measurement ranking could be performed on the basis of a cheaper screening interview and the technique of RSS may prove to be a timely innovation.
- (v) With the availability of computerized Geographic Information Systems (GIS), ranking prospective sample locations across a landscape may be

done rapidly prior to expensive field visits, thus allowing RSS to be applied to large scale surveys to obtain a more precise estimate at reduced cost. If prospective locations are selected at random from across a region and allocated to a set, then each location can be referenced to data layers in a GIS and, based on a derived ranking index, each member of the set can be ranked relative to each other. This merger of GIS and RSS has been recommended by Johnson and Myers (1993) and by Myers, Johnson and Patil (1994).

- (vi) Following a catastrophic event such as flooding or fire, those in charge of management and planning of natural or cultural resources need rapid assessments of the spatial extent and magnitude of damage. The authors cited above in item (v) recommend that the combination of RSS and Geographic Information Systems (GIS) can result in rapid mobilization of available information to design a very efficient field sampling strategy.

4. Summary

Compared to simple random sampling, the ranked set sampling method has been proven theoretically and shown empirically to yield more precise estimators of the population mean. This is especially desirable when sample sizes are generally small as with environmental data which are expensive or destructive to obtain. A browse through the references cited throughout this publication will also reveal that many other population features can be estimated with higher precision using ranked set sampling.

So long as ranking is not cost-prohibitive, such as with exploiting available expert judgment or using a readily available ranking covariate, ranked set sampling can serve to achieve observational economy. Even if measurable effort is required to obtain values of a covariate, it may still be worth while if the resulting rankings were reasonably accurate. And this is particularly so, because there is considerable robustness in the ranked set sampling procedure.

A very attractive feature of ranked set sampling is that, unlike other double sampling procedures, it can use subjective expert opinion as the source of auxiliary information. Such a feature also appeals to the philosophy of total quality management because it exploits the expertise of not only a professional statistician, but also of field personnel who usually know the most about the population being sampled.

References

- COBBY, J. M., RIDOUT, M. S., BASSETT, P. J., AND LARGE, R. V. (1985). An investigation into the use of ranked set sampling on grass and grass-clover swards. *Grass and Forage Science*, **40**, 257-263.
- DAVID, H. A. AND LEVINE, D. N. (1972). Ranked set sampling in the presence of judgment error. *Biometrics*, **28**, 553-555.
- DELL, T. R. (1969). The theory and some applications of ranked set sampling. Ph.D. Thesis, Department of Statistics, University of Georgia, Athens, Georgia.
- DELL, T. R. AND CLUTTER, J. L. (1972). Ranked set sampling theory with order statistics background. *Biometrics*, **28**, 545-553.
- EVANS, M. J. (1967). Application of ranked set sampling to regeneration surveys in areas direct-seeded to longleaf pine. Masters Thesis, School of Forestry and Wildlife Management, Louisiana State University, Baton Rouge.
- HALLS, L. K. AND DELL, T. R. (1966). Trial of ranked set sampling for forage yields. *Forest Science*, **12**, 22-26.
- JEWISS, O. R. (1981). Shoot development and number. In *Sward Measurement Handbook*, J. Hodgson, *et al.*, eds. Hurley: The British Grassland Society. pp. 93-114.
- JOHNSON, G. D., AND MYERS, W. L. (1993). Potential of ranked-set sampling for disaster assessment. Presented at IUFRO S4.02 Conference on "Inventory and Management Techniques in the Context of Catastrophic Events," June 1993.
- KAUR, A., PATIL, G. P., AND TAILLIE, C. (1994). Unequal allocation model for ranked set sampling with skew distributions. Technical Report 94-0930, Center for Statistical Ecology and Environmental Statistics, Department of Statistics, Pennsylvania State University, University Park, PA.
- KAUR, A., PATIL, G. P., SINHA, A. K. AND TAILLIE, C. (1995). Ranked set sampling: An annotated bibliography. *Environmental and Ecological Statistics*, **2**(1) (to appear).
- MARTIN, W. L., SHARIK, T. L., ODERWALD, R. G., AND SMITH, D. W. (1980). Evaluation of ranked set sampling for estimating shrub phytomass in Appalachian oak forests. Publication Number FWS-4-80, School of Forestry and Wildlife Resources, Virginia Polytechnic Institute and State University, Blacksburg, Virginia.
- MCINTYRE, G. A. (1952). A method for unbiased selective sampling, using ranked sets. *Australian Journal of Agricultural Research*, **3**, 385-390.
- MYERS, W., JOHNSON, G. D., AND PATIL, G. P. (1994). Rapid mobilization of spatial/temporal information in the context of natural catastrophes. In

- 1994 *Proceedings of the Section on Statistical Graphics*, American Statistical Association, Alexandria, VA. pp. 25-31.
- MUTTLAK, H.A. AND McDONALD, L.L. (1992). Ranked set sampling and the line-intercept method: a more efficient procedure *Biom. J.*, **3**, 329-346.
- PATIL, G. P., SINHA, A. K., AND TAILLIE, C. (1994). Ranked set sampling. In *Handbook of Statistics, Volume 12: Environmental Statistics*, G. P. Patil and C. R. Rao, eds. North Holland/Elsevier Science Publishers.
- STOKES, S. L. (1977). Ranked set sampling with concomitant variables. *Communications in Statistics—Theory and Methods*, **A6**, 1207-1211.
- STOKES, S. L. (1980a). Estimation of variance using judgment ordered ranked set samples. *Biometrics*, **36**, 35-42.
- STOKES, S. L. (1980b). Inferences on the correlation coefficient in bivariate normal populations from ranked set samples. *Journal of the American Statistical Association*, **75**, 989-995.
- STOKES, S. L. AND SAGER, T. W. (1988). Characterization of a ranked set sample with application to estimating distribution functions. *Journal of the American Statistical Association*, **83**, 374-381.
- TAKAHASI, K. AND WAKIMOTO, K. (1968). On unbiased estimates of the population mean based on the sample stratified by means of ordering. *Annals of the Institute of Statistical Mathematics*, **20**, 1-31.
- YANAGAWA, T. AND S.H. CHEN. (1980). The MG procedure in ranked set sampling. *J. Statist. Plann. Inference*, **4**, 33-34.